COGNITIVE SCIENCE 14, 253–280 (1990)

Detection of Cognitive Structure with Protocol Data: Predicting Performance on Physics Transfer Problems

WILLIAM C. ROBERTSON
The Biological Sciences Curriculum Study

This article presents a cognitive map proposed to be associated with understanding of the "system concept," one component of the physics principle of Newton's second law. A definition of the concept is followed by the results of a problem-solving experiment designed to investigate whether or not good problem solvers possess cognitive structures similar to the one proposed. Think-aloud protocols were collected as subjects solved a series of physics problems involving Newton's second law. Coding schemes were used to analyze these protocols and to develop a quantitative index intended to reflect the extent to which subjects possessed certain components of the proposed structure. This index proved a highly significant predictor of performance on Newton's second law transfer problems contained in a written exam. In contrast, performance on a set of familiar problems in the same written exam was not a good predictor of performance on transfer problems. These results indicate that cognitive structures that connect the equation $\Sigma F = ma$ and the concept of choosing appropriate systems are at least part of what constitutes understanding of the principle of Newton's second law. Implications for physics education and problem-solving research are presented.

Transfer problems—problems that are structurally, but not conceptually unfamiliar to the solver—have long been used as a measure of understanding (Egan & Greeno, 1973; Gagné & Brown, 1961; Katona, 1940; Perkins & Salomon, 1989; Royer, 1986; Wertheimer, 1945). Students who rely on memorized algorithms for solving problems typically do not perform as well on transfer problems as do students who rely on an understanding of the underlying concepts. A clear picture of the cognitive structures that support conceptual understanding does not, however, exist. The theoretical models of Ausubel (1968) and Gagné and White (1978) suggest that "connections"

The author gratefully acknowledges the conceptual and methodological guidance of K. Anders Ericsson, and the thoughtful and extensive reviews of this article by Andrea A. diSessa and Jill H. Larkin.

Correspondence and requests for reprints should be sent to William C. Robertson, The Biological Sciences Curriculum Study, 830 North Tejon, Suite 405, Colorado Springs, CO 80903.

among propositions, images, episodes, and intellectual skills within a person's memory structure promote understanding and enhance performance on transfer problems. These models also imply that, to understand new concepts, one must anchor the new concepts to existing structures.

In support of the theoretical models, Mayer (1974) and Mayer, Stiehl, and Greeno (1975) have found that instructional materials that tie unfamiliar concepts to everyday experiences help students perform well on transfer problems. Presumably, students who are exposed to this type of instruction have established the cognitive connections necessary for understanding. With notable exceptions (Bromage & Mayer, 1981; Moreira & Santos, 1981), however, there is little evidence that various instructional materials actually help students acquire the cognitive structures suggested by the theoretical models.

In the work reported here, analysis of data from verbal reports, or think-aloud protocols, was used to investigate the kinds of cognitive structures that are associated with successful performance on transfer problems. Demonstrations that protocol analysis can produce detailed, quantifiable, and reproducible data on human thought processes (Ericsson & Simon, 1984) suggest its use in this regard.

The problem-solving domain investigated in this study is that of Newton's second law—a major principle in the physics of mechanical phenomena. Physics is a notoriously difficult subject, and common sense interpretations of principles are most often insufficient for understanding. Therefore, one can expect considerable variation in novice students' understanding of principles such as Newton's second law, and hence their success in solving transfer problems requiring use of the principles. The following section outlines what constitutes understanding of one component of the principle of Newton's second law, the "system concept," which served as the focus of this investigation.

NEWTON'S SECOND LAW AND THE SYSTEM CONCEPT

Newton's second law states that the sum of the external forces acting on a system, signified by ΣF, is equal to the mass of that system multiplied by the acceleration of the system. In equation form: $\Sigma F = ma$. As with many physics principles, a conceptual understanding of Newton's second law should involve connecting this seemingly simple equation with its attendant ancillary knowledge. Far from being subordinate to the equation, this ancillary knowledge gives meaning to the equation, making it the representation of a principle, rather than a formula into which one plugs various numbers. Eylon and Reif (1984) and Heller and Reif (1984) suggest that an expert's organization of this knowledge is structured hierarchically, starting with broad, general principles at the top of the hierarchy and proceeding to sub-

sidiary concepts and details at lower levels. That notion is supported by Chi, Feltovich, and Glaser (1981). The ancillary knowledge contained in the expert's hierarchical organization includes qualitative and quantitative descriptions of the principle (Reif, 1985) as well as procedural specifications of the principle and a knowledge of the conditions under which the principle is applicable (Larkin, 1981).

For Newton's second law, the ancillary knowledge includes at least the following:

1. An understanding of the concepts of force, mass and acceleration;
2. Knowledge that $\Sigma F = ma$ is a vector equation, and that the direction of the acceleration must be the same as the direction of the net force;
3. Knowledge that forces cause accelerations, and not vice-versa;
4. Knowledge that the use of the equation involves the choice of a system to which the equation will be applied;
5. Knowledge that the choice of system has definite implications for the application of the equation;
6. Knowledge of the conditions under which using the equation is likely to give useful information.

This list focuses on what White (1985) refers to as "internal associations"—connections among the essential parts of a principle that, along with external connections such as between the principle and everyday experiences, facilitate understanding.

The present study deals with one part of the listed ancillary knowledge: the concept of strategically choosing systems in order to relate known and desired information. A *system* in relation to $\Sigma F = ma$ refers to an object, a portion of an object, or a collection of objects that one wishes to consider. For the purpose of this article, the "system concept" will refer not just to the isolation of a set of objects, but doing so in a way meaningful for the application of Newton's second law. Figure 1 shows a concept map (Novak & Gowin, 1984) that details a proposed expert's structure of the system concept. This map represents a minimum set of connections that would enable a problem solver to choose appropriate systems and take actions consistent with those choices, in solving a wide variety of Newton's second law problems. Other important aspects of solving the problem, such as deciding what factors to include in ΣF, are not represented. The map is presented here as a clarification of the system concept and as a guide for understanding the protocol analysis to be described later.

Referring to Figure 1, a choice of system involves more than a random selection of objects, and acquires meaning, when the problem solver uses the following criteria for the choice. One chooses a system because one wants to know something about (a) the external forces acting on the system, (b) the mass of the system, or (c) the acceleration of the system. Because the

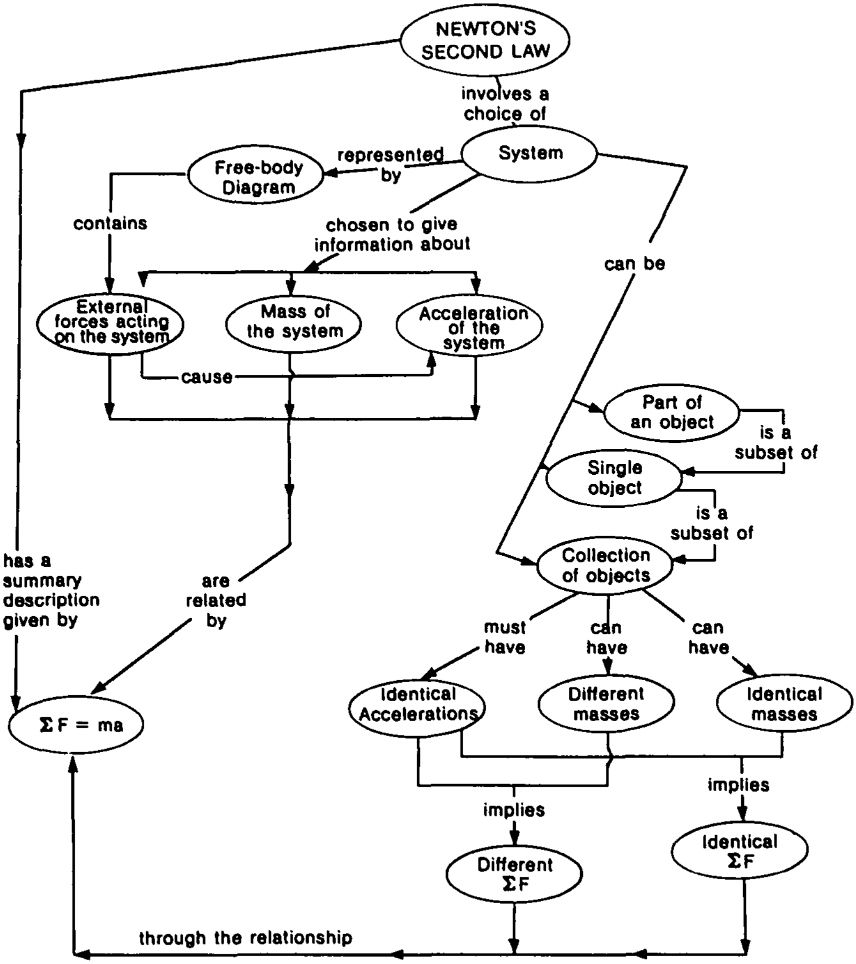


Figure 1. Proposed expert's structure of the system concept.

a in $\Sigma F = ma$ refers to the acceleration of the center of mass of the system (Sherwood, 1983), any objects in a collection that is considered as a system must have identical accelerations. Otherwise, the acceleration obtained from applying $\Sigma F = ma$ will, in general, not refer to the acceleration of any one of the objects, and would thus not describe a useful quantity in problems typically encountered in an introductory course. The restriction of identical accelerations, coupled with a knowledge of whether the objects in the collection have different or identical masses, implies something of the nature of the forces acting on the individual objects (see Figure 1). For example, different masses and identical accelerations of two objects implies that different sums of external forces are acting on the objects.

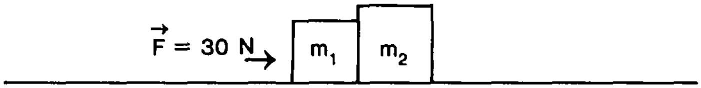


Figure 2. A force of 30 N is applied continuously to a box of mass $m_1 = 10$ kg. This box is in contact with a box of mass $m_2 = 20$ kg, and both boxes are on a frictionless table. Assuming the boxes remain in contact, find the acceleration of m_2.

To illustrate the system concept further, consider the problem shown in Figure 2. There are two possible paths to the solution of this problem; each path involves a different choice of systems. One can either consider each mass separately or the two masses together. Figure 3 shows a task analysis for the problem. Here the term, *task analysis,* refers to the steps a problem solver is likely to use in following one of the paths to a correct solution to the problem. The task analysis for this problem, as well as for all the problems used in this study, was derived from an informal analysis of think-aloud protocols collected during pilot studies. The solution paths shown are not the only ones possible (for example, energy methods will work), but they should show the reader typical steps taken by problem solvers en route to a solution, as well as steps in the solution path that involve a conscious or unconscious choice of systems. The starred boxes in the task analysis indicate operations related to a subject's understanding of the system concept.

As is evident from the two paths in the task analysis, the choice of system determines the appropriate quantities to substitute for ΣF and for m in the force equation. For example, if one is applying Newton's second law to m_2 alone, it would be inappropriate to use a value of 30 Newtons for ΣF, because 30 Newtons is the proper ΣF for the system of both masses, considered together. It is also possible to choose a system that will yield no useful results. For example, choosing the blocks and the table they are resting on as a system is a loss, because the blocks-table system has components with different accelerations (the a in $\Sigma F = ma$ would then refer to the center of mass of the system and not to the desired acceleration of m_2) and one component (the table) with an unknown mass.

OVERVIEW OF THE PROCEDURE

In this study, subjects thought aloud while solving a series of Newton's second law problems. The resulting protocols were then coded, the codes being used to determine the extent to which subjects did or did not possess cognitive structures similar to that represented in Figure 1. This indication of subjects' knowledge of the system concept was then correlated with performance on transfer problems. A high correlation between the protocol-derived indicator and performance on transfer problems would validate the notion that the proposed cognitive structure is associated with a conceptual understanding of Newton's second law.

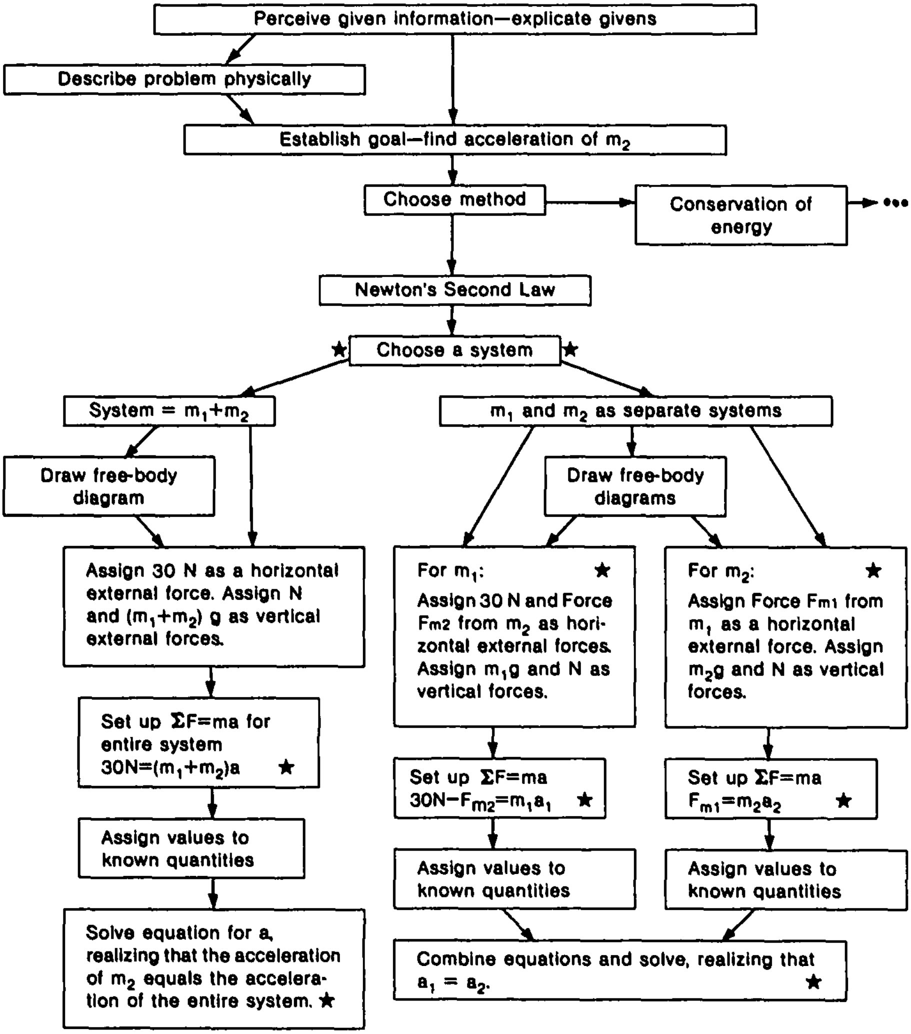


Figure 3. Task analysis for the problem in Figure 2.

PROBLEM-SOLVING EXPERIMENT

Method

Subjects and Design. The 20 subjects were paid volunteers who were approximately two-thirds of the way through their first semester of non-calculus-based college physics. All the subjects were receiving a C or better in physics at the time of their selection. Data was recorded for all subjects,

but 15 were randomly chosen for purposes of analysis. All the subjects took the same series of tests, so comparisons were within-subject or correlational.

Materials. Materials consisted of three physics problems used in the tape-recorded, think-aloud sessions and a written test consisting of seven physics problems. These problems are in the Appendix.

Three Newton's second law problems were used in the think-aloud sessions. These problems all involve two masses (Problem 3 involves three masses, but one is stationary and doesn't figure into the solution of the problem). A correct solution for each of these problems requires choosing *a* system, consciously or unconsciously. There are few incorrect system choices—most choices have the possibility of leading to a correct solution. Because the subjects had encountered these three problems (or ones like them) before in their physics courses, they should have had an idea of how to proceed. The crucial aspect of these problems is that one can make a system choice (possibly because his or her instructor or textbook made that choice in a similar problem in a physics class) without having an understanding of the system concept. But because system choices are involved, the probability exists that students who understand the system concept will verbalize statements relative to that concept. The problems thus provide a medium for differentiating between subjects who do and do not understand the system concept.

The written test consisted of an instruction page, four familiar problems, and three transfer problems. The familiar problems were similar to problems the subjects might have encountered in their lectures, homework, and exams. The transfer problems required more than one choice of system for correct solution, and there were many system choices that would not lead to a correct solution. Because the transfer problems involved nonstandard system choices, algorithms that might have led to successful performance on more familiar problems would be of no use. It was thus unlikely that someone who did not understand the system concept could do well on the transfer problems unless they had been given an algorithm that specifically applied to these problems. The transfer problems did not involve frictional forces, did not involve vector addition or resolution of vectors into components, and did not involve the concept of Newton's third law (action-reaction). As such, they represented transfer problems only with respect to the system concept. A transfer problem and its solution are illustrated in Figure 4. Familiar and transfer problems appeared alternately on the test. A sheet of relevant and irrelevant formulas was available for the subjects to use throughout the testing.

Procedure. In individual, tape-recorded, one-half to three-quarter hour sessions, the subjects were instructed to think aloud as they attempted to solve the three protocol problems. The experimenter gave each subject warm-

Blocks 1, 2, 3, 4, and 5 each have a mass of 5 kg, and are connected by ropes of negligible mass. A force of 400 Newtons is applied horizontally to block 5. The blocks are sitting on a frictionless surface.

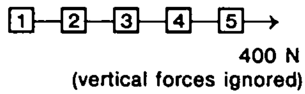
a) Find the acceleration of block 5.

b) Find the tension of the rope between blocks 3 and 4.

Solution to part (a)

The easiest way to do this (there are other ways) is to consider all 5 blocks as one system:

$$\Sigma F = ma$$
$$400 \text{ N} = (m_1+m_2+m_3+m_4+m_5)a$$
$$a = 400 \text{ N}/25 \text{ kg}$$
$$= 16 \text{ m/s}^2$$



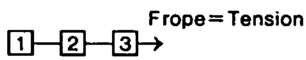
The acceleration of block 5 is the same as the acceleration of the entire system.

Solution to part (b)

The problem asks for the tension, or the force the rope exerts on block 3 (also on block 4). Thus we *must* choose a system such that this force is an external force acting on that system. Two possible solutions are shown below.

System = blocks 1, 2, and 3:

$$\Sigma F = ma$$
$$\text{Tension} = (m_1+m_2+m_3)a$$
$$= (15 \text{ kg})(16 \text{ m/s}^2)$$
$$= 240 \text{ Newtons}$$



System = blocks 4 and 5:

$$\Sigma F = ma$$
$$400\text{N} - \text{Tension} = (m_4+m_5)a$$
$$\text{Tension} = 400\text{N}-(m_4+m_5)a$$
$$= 400\text{N}-(10\text{kg})(16\text{m/s}^2)$$
$$= 240 \text{ Newtons}$$

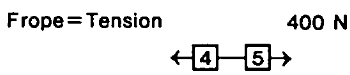


Figure 4. A transfer problem and its solution.

up exercises until it was believed the subject was relatively comfortable with the process of thinking aloud. Subjects could write down anything they wished as well as use a calculator and formula sheet.

From 7 to 23 days after the think-aloud session, the subjects took the written exam. They had as much time as they needed. The exams were graded by two graduate students in physics, who devised their own system of assigning points. Both graduate students graded three of the written problems, resulting in an intergrader correlation of .87.

Analysis. Tape-recordings of subject protocols were transcribed verbatim. These transcriptions were then separated into statements, which were defined as any phrases separated by pauses in speech, with the criterion that phrases

would be combined if they were grammatically or logically connected. Each statement was then assigned a code based on extensive coding schemes developed by the experimenter.

Heller & Reif's (1982) specifications of the general procedures necessary for efficient problem solving provided the broad categories for the coding schemes. As such, each coding scheme contained six categories:

1. Basic Description of the problem
2. Theoretical Description of the problem
3. Exploratory Analysis
4. Metacognitive Statements
5. Problem Solution
6. Assessment.

Task analyses such as the one in Figure 3 were then used as a guide for developing codes within the major categories, that were specific to each of the three problems used in gathering protocols. Protocols gathered in pilot studies led to further modification of these problem-specific codes. This modification consisted of adding codes to account for unanticipated verbalizations, and refinement of others to reflect fine distinctions between different verbalizations.

The coding of statements was accomplished with the aid of a computer program, MPAS (Ericsson & Simon, 1984), that presented the statements one at a time to the coder and asked for a code for each statement. MPAS has the capability to choose a "problem 1 statement" randomly from among the data for all 15 subjects, so that the coder was aware of which *problem* was being coded, but not which *subject* was being coded. Coding with MPAS was not done entirely without contextual cues, as statements immediately preceding, and immediately following, could be displayed by the coder. The coding of statements was performed by the experimenter, but a high correlation (.90) between codes assigned by the experimenter, and codes assigned by a "naive" coder, attest to the fact that the assigned codes are replicable. The naive coder had little knowledge of physics or of the research project.

In the experiment reported here, 2,396 protocol statements were analyzed and 2,191 of these, or 91.4%, were assigned codes from the coding schemes. The remaining statements were labeled NC (no code) and were not included in the analysis of data for the experiment.

Codes that described processes in the starred portions of the task analysis (Figure 3), or described in the connections illustrated in Figure 1, were considered as possible indicators that subjects did or did not understand the system concept. These codes represent approximately 40% of the total number of codes in the coding scheme, and were assigned only to about 30% of the statements in the subject protocols. Thus, the majority of the protocol statements reflected problem-solving processes not related to the system concept.

Codes that were associated with the system concept were separated into positive (understand) and negative (don't understand) indicators. Categories for these indicators are presented below. Following each category are examples taken from subject protocols of statements that were assigned codes belonging to that category. Unless otherwise noted, the examples are from Problem 1 protocols. A list of codes used in the analysis of the first protocol problem, as well as two sample, coded protocols, are in the Appendix.

Positive Indicators

1. Explicit Choice of a System. A statement of choice reflects an awareness of the appropriate connection in Figure 1. Note that the choice must be explicit and not implied through other actions by the subject. Examples from protocols:

> But. . .then the second part, you have to separate out m_2, its mass, and find out. . .the forces. . .

> Ok. . .assuming the boxes remain in contact. . .you can assume they act as. . . together. . .as one. . .

2. Explicit Statement that a Choice of System is Possible. Such a statement goes further than statements in Category 1 in that the subject not only makes a choice, but indicates that the choice is not the only one possible. Figure 1 illustrates the possible choices and Figure 3 shows the junction in the problem-solving process where this might occur. Examples from protocols:

> It could either be figured out individually. . .for each of the two, or it could pull all together. . .and done as one. . .(Problem 2)

> So. . .do I want to separate the block and the boy? (Problem 3)

3. Statements that the Accelerations of the Two Masses in the Problem are the Same. Because each protocol problem asks for the acceleration of *one* of the two objects involved, a reference to the acceleration of the second object, or to both objects together, might indicate a problem-solving strategy involving the choice of an alternate system (see Figure 1). It might also indicate a realization that the components of a system must have identical accelerations, as is also indicated in Figure 1. Confounding this indicator, though, is the possibility that a subject could make this kind of statement based solely on a physical and theoretical description of the problem. Examples from protocols:

> And so that's how I found the acceleration of m_2, it had to be the same. . .as the acceleration of the whole. . .masses, since they were. . .in contact. . .

> And acceleration's going to be the same...so I'll do an indirect way of doing it...I'll find the acceleration of m_2...acceleration of m_2...which would be equal to...m_1...plus m_2, which is the total mass... (Problem 2)

4. Statements that the Forces Acting on the Two Masses Must be Different. All of the protocol problems involve two masses with identical accelerations, so the implied conclusion (see the bottom right of Figure 1) is that the forces acting on the masses must be different. This conclusion is counterintuitive (see Category 6, and Results and Discussion), and unlikely to be made on the basis of a purely physical or theoretical description. Examples from protocols:

> So I would say...that...the force to accelerate mass two...plus the force to accelerate mass one...plus force one equals the force to accelerate the whole system...

> And that the total...force...would equal 60, but the force that it needed to move the boy and block would have to be more than the block, since the masses were different...and so it would be...dispersed that way... (Problem 3)

5. Substitution of Algebraic Symbols or Numbers into $\Sigma F = ma$ that are Consistent with a Given System Choice (Including Stating that Given Forces are Acting on an Appropriate System). The system concept provides strong constraints on relations between the algebraic symbols used and their referents, as indicated by the qualifiers in "mass *of the system*," "acceleration *of the system*," and "forces acting *on the system*." Although obeying these constraints does not guarantee that the problem solver did so by virtue of understanding the system concept, it can be argued that correctly answering protocol Problems 1b and 3 without an understanding of the concept would be unlikely. Although this indicator was included in the analysis, the results of the experiment are significant, though less so, when this indicator is removed from the analysis (see Table 3). Examples from protocols:

> m_2 hanging down...m_2g minus...tension pulling up...equals m_2a...ok... so now, tension...is equal to...the force to accelerate m_1...which is m_1a... (Problem 2)

> So the force is applied to...the two masses, m_1 plus m_2...

Negative Indicators

6. Statements that the Forces Acting on the Two Masses are the Same. Such a statement suggests that the connections illustrated in the bottom right of Figure 1 are missing. There are indications, however, that statements in this category are highly intuitive and the subjects often made them without

any apparent need for justification (see Results and Discussion). Examples from protocols:

> There was a force of 30 Newtons applied to both of them...but...that was still applied to m_2...

> So let's say that there's 30...if it's being applied to m_1, it's got to be applied to m_2...at the same time...

7. Statements that the Accelerations of the Two Masses are Different. These statements are indicative of mindlessly applying the force equation to the two masses, and are usually the result of assuming that the forces acting on the two masses are the same (Category 6). Again refer to the bottom right of Figure 1. Example from protocols:

> I would think that the acceleration of mass two...would be half the acceleration of mass one...because it's twice as much...weight as mass one.

8. Inappropriate Substitution of Symbols or Numbers into the Force Equation. In contrast to Category 5 above, such statements indicate a *lack* of understanding of how the symbols or numbers one uses for ΣF, m, and a relate to a given choice of system. Example from protocols:

> So that's the force...the force acting on m_2 is...the mass times the acceleration...of m_1...

An "index of system-concept understanding" was assigned to each subject for each problem. This index of understanding was based upon the presence of positive or negative indicators of system-concept understanding and was a numerical sum of positive and negative indicators, correcting for duplication, for each problem.

Physics grades were obtained from the instructors of the courses in which the subjects were enrolled, and ACT math and verbal scores were obtained from the college registrar. Five of the subjects had only SAT math and verbal scores, so ACT equivalent scores were calculated for those subjects using a table constructed by Langston and Watkins (1980) and provided by the Office of School and College Relations at the University of Illinois.

A multiple regression was performed (using SPSSX) with performance on transfer problems as the dependent variable. The following variables were used as predictors: the index of system-concept understanding, a sum of ACT math and verbal scores, a sum of scores on the physics course final exam and the physics course exam covering the principle of Newton's second law, performance on familiar problems, and sex.

TABLE 1
Correlation Matrix Showing the Relationships Between Transfer Performance and Predictor Variables

	Transfer Problems	System-Concept Index	ACT Scores	Physics Exams	Sex
System-concept Index	.762**				
ACT scores	−.343	−.514*			
Physics exams	−.166	.032	.358		
Sex	.045	−.007	−.063	.207	
Familiar problems	.015	−.095	.393	.409	−.321

* $p<.05$, ** $p<.001$

TABLE 2
Results of Multiple Regression to Predict Transfer Performance[1]

Variable in Equation[2]	F	Significance of F	Part Correlation	Square of Part Correlation
System-concept index	16.00	.002	.722	.521
Physics exams	1.75	.212	−.239	.057
ACT scores	0.74	.408	.155	.024
				Sum .602

[1] Dependent variable: Performance on Transfer Problems Multiple r squared=.642
[2] Regression Equation:
Predicted transfer performance = 9.76 × (system-concept index) − 0.03 × (physics exams) + 0.39 × (ACT scores) − 23.50

RESULTS AND DISCUSSION

Predicting Performance on Transfer Problems

A correlation matrix of relevant variables is shown in Table 1. The index of system-concept understanding is the only significant single predictor of performance on transfer problems. The high, negative correlation between the index of system-concept understanding and ACT scores appears to be an artifact of two outliers in the data. These two subjects had the lowest ACT scores and were among the highest scorers on transfer performance and the index of system-concept understanding. The correlations of sex and performance on familiar problems with performance on transfer problems are so low that these variables were not used in the regression analysis.

The results of the regression analysis for predicting transfer performance are shown in Table 2. The square of the part correlation coefficient for each variable tells how much of the variance in transfer performance can be

uniquely explained by that variable, once all other variables are in the regression equation. The fact that these unique contributions added together approximately equal the multiple r squared of .642 indicates that there is little shared variance among the variables and that they are essentially independent predictors of performance on transfer problems. Considering *only the amount of variance in transfer performance explainable by these three variables,* 4% of this variance can be attributed to ACT scores, 9.5% can be attributed to scores on physics course exams, and 86.5% can be attributed to the index of system-concept understanding. Thus, it appears that understanding of the system concept is an important component of the ability to solve Newton's second law problems of the type used as transfer problems in this study.

Reference to the interaction shown in Figure 5 can make the above results more transparent. Subjects were grouped according to low (−1–0), medium (1), and high (2–3) scores on the index of system-concept understanding. There is no significant difference between the low, medium, and high groups on scores on the familiar problems, but the high group performed significantly better than the low group ($p < .05$) on the transfer problems. The low-index group (poor understanding of the system concept) showed marked differences in performance between familiar and transfer problems. The medium group had less of a difference in performance, and the high group performed equally well on familiar and transfer problems. An interaction such as this is common when one compares groups that use understanding and groups that use memorization in problem solving (Egan & Greeno, 1973; Mayer, 1974; Mayer et al., 1975), and this interaction strengthens the argument that the index developed here is actually associated with understanding of a concept.

Although the index of system-concept understanding is correlated with performance on transfer problems (and hence what is commonly termed "understanding"), and was designed to reflect the existence of cognitive structures similar to those in Figure 1, there may be some debate as to what the index is actually revealing. As was noted in the explanations of some of the positive and negative indicators, subjects' physical and theoretical problem descriptions might contribute to a high score on the index. In addition, one must be cautious about claiming anything like a one-to-one correspondence between subject verbalizations and specific nodes and links in a proposed cognitive map.

A subject's knowledge about the behavior of physical systems is not limited to Newton's second law, so a complete cognitive map would be more extensive than Figure 1 indicates. It is as though a map that concentrates on one part of one principle is but a slice of a complete, multidimensional map (West, Fensham, & Garrard, 1985). Therefore, the reader will undoubtedly be able to propose alternative interpretations of the indicators used here.

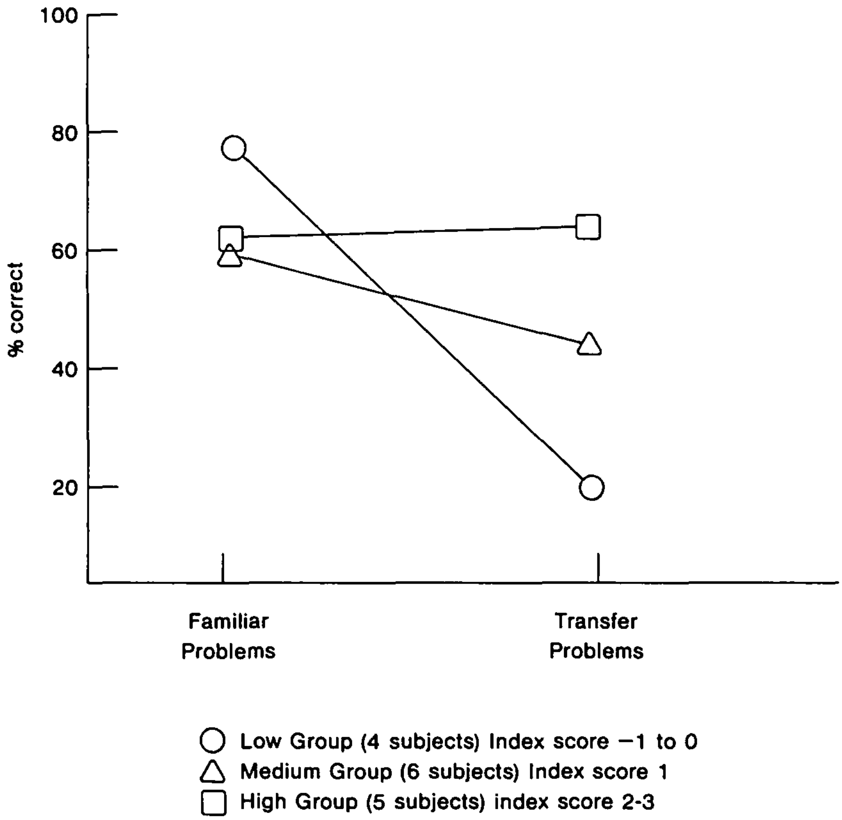


Figure 5. Interaction between understanding of system concept (high, medium, and low index scores) and scores on familiar and transfer problems.

While acknowledging the resulting ambiguity this creates for interpreting the results of this study, one can only wait for more complete cognitive models to resolve the ambiguity.

It will help, however, to point out two things that are *not* associated with the index. First, it does not seem to be measuring a general problem-solving ability, at least not problem-solving ability as measured by standardized tests and typical course exams. The index is a very poor predictor of performance on the familiar problems contained in the written exam, and it is not significantly correlated with performance on physics course exams. Also, the index of system-concept understanding adds significantly more information for predicting transfer performance, than a score that simply reflects the number of protocol problems subjects solved correctly. This is illustrated with a regression analysis in Table 3. The score on protocol problems

TABLE 3
Results of Multiple Regression Comparing the Index of System-Concept Understanding and Scores on Protocol Problems as Predictors of Transfer Performance[1]

Variable in Equation[2]	F	Significance of F	Part Correlation	Square of Part Correlation
Score on protocol problems (range 0 to 6)	0.60	.453	−.141	.020
System-concept index	13.15	.003	.661	.437
				Sum .457

[1] Dependent variable: Performance on Transfer Problems Multiple r squared=.775
[2] Regression Equation:
Predicted transfer performance=9.93×(system-concept index)−3.72×(score on protocol problems)−28.9

is, of course, highly correlated with the index of system-concept understanding, and was not used in the previous regression (Table 1). Second, the index does not seem to be measuring any sort of ability to verbalize thought processes. There is no significant difference in the lengths of protocols between people with high- and low-index scores. In fact, some of the subjects with the highest index scores had protocols of the shortest lengths.

It does seem clear that knowledge structures that are similar to those represented in Figure 1 are at least part of what constitutes understanding of Newton's second law. Such a finding supports theoretical claims that the definition of understanding be extended from performance on transfer problems to include the existence of these or related structures. However, the limited nature of the problems used in this study makes generalization of results to other types of Newton's second law problems an uncertain task. The effects of including complicating factors such as vector resolution and friction are unknown, but it is obvious from the results of pilot studies that difficulties with these factors can disrupt a subject's problem-solving process enough that concern with more subtle matters such as the system concept is completely masked.

Analysis of Errors

It is possible that the codes assigned to subject protocols are sufficiently detailed that one could predict from these codes not only subject performance on transfer problems, but also the types of errors that subjects committed on familiar and transfer problems. Although the attempt to use protocol statements to predict types of errors proved inconclusive, the presence of one particular type of error was enlightening. Aside from errors associated with choosing systems, the most common error committed by subjects on both the protocol and written problems was that of assuming that the force

acting on one object in a group of connected objects was transmitted—undiminished—to other objects in the group. All but 2 of the 15 subjects committed this "contact force" error at least once. Thus, even subjects who scored well on the system-concept index committed the error. That good problem solvers made a contact force error is not surprising in light of an informal study by Arons (1982), in which it was found that 13 of 17 entering physics graduate students made the same type of error when confronted with a problem similar to protocol problem Number 2.

Why is such an error so common? It appears that it is not the result of a misunderstanding of the system concept in all cases, but of a reliance upon (incorrect) intuition to make a decision, instead of thinking through the implications of that decision in the context of Newton's second law. This conclusion is supported by the fact that many subjects made an appropriate system choice on the transfer problems, but then relied upon intuition to solve the problem rather than follow through the implications of the correct system choice in applying Newton's second law. It should be noted, however, that there was a high correlation ($r = .585$, $p = .001$) between the frequency of contact force errors in subject protocols and the frequency of explicit errors associated with choosing systems. Thus, the two issues are not completely independent.

IMPLICATIONS FOR EDUCATION

That the system concept is central to understanding Newton's second law should be no surprise to physics educators. As is evident from their performance on the transfer problems in this study, however, the students in a typical noncalculus-based introductory physics course do not have a clear understanding of the system concept. One could easily dismiss this result by noting that novice physics students do not understand many basic concepts, including acceleration and momentum. A more productive approach, though, would be to ask how present instruction might be altered to bring about understanding of the concept. Through pedagogy, that is not the subject of this article, the students should be made aware of the connections among the various components of the system concept and its relationship to general physics principles.

Although the currently popular notion of tying physics principles to everyday situations (establishing external connections) is a promising strategy for promoting understanding, it should be clear from the results of the present study that connections also need to be made *among* scientific concepts (establishing internal connections), even if those connections are sometimes *divorced* from life's everyday situations. The issue of contact force errors is a good example of this. If one has a series of objects in contact and they are not accelerating (as is usually the case in everyday experience), then a force

applied to one object *is* transmitted undiminished to the other objects. It is only when these objects are accelerating (not a common occurrence) that the force is not transmitted undiminished. In this situation, which is common for physics problems but uncommon for the real world of the average person, the internal consistency of physics concepts must help guide the student's thinking, along with connections between the problem and everyday experience.

IMPLICATIONS FOR PROBLEM-SOLVING RESEARCH

The usefulness of protocol analysis for gaining qualitative insights into cognitive processes is well accepted, but the work presented here indicates that think-aloud protocols can be used quantitatively. It is entirely possible to generate a replicable set of data (assigned codes) that one may use to make specific predictions about problem-solving performance. The process of generating this replicable data set should be strongly grounded in a theoretical model of the organization of relevant knowledge, such as a central principle and its attendant ancillary knowledge, as well as models for the problem-solving processes subjects will likely go through in the given task.

This process is specific to individual problems. Generic task analyses and resulting coding schemes, which are designed to apply to a variety of problems, cannot hope to capture the details necessary to probe connections among concepts that manifest themselves differently in different problem-solving situations. This is especially true in light of the vast amount of research that demonstrates the importance of domain-specific knowledge as opposed to general skills or heuristics in problem-solving tasks.

This study also shows that it is useful to study problem-solving procedures by looking at subjects who are all at the same level of instruction. There appears to be sufficient variation in both problem-solving performance and methods of problem solving to allow a researcher to gain insights into the knowledge structures and procedures that facilitate successful performance. Thus, investigation of subjects at the same level of instruction provides a viable alternative to expert–novice studies, in which it is difficult to draw conclusions based upon performance of subjects who differ considerably in knowledge base and problem-solving procedures.

Finally, it should be clear that this study provides only a small clue for the researcher who wishes to know "what is possessed" by people who utilize conceptual understanding in a given domain. The presence of indicators that compose the index of understanding presented here is *sufficient* for predicting transfer performance, but are all of the indicators *necessary*? Would a more complete cognitive model allow one to establish more of a one-to-one relationship between the model and protocol statements? It would be pleasing to see more studies that directly address these questions rather

than rely upon implications from various instructional materials. Such research should produce detailed task analyses of problems being studied and relate performance on these problems to the task analyses and the cognitive structures that they imply.

REFERENCES

Arons, A.B. (1982). Phenomenology and logical reasoning in introductory physics courses. *American Journal of Physics, 50,* 13–20.

Ausubel, D.P. (1968). *Educational psychology: A cognitive view.* New York: Academic.

Bromage, B.K., & Mayer, R.E. (1981). Relationship between what is remembered and creative problem-solving performance in science learning. *Journal of Educational Psychology, 73,* 451–461.

Chi, M.T.H., Feltovich, P., & Glaser, R. (1981). Categorization and representation of physics problems by experts and novices. *Cognitive Science, 5,* 121–152.

Egan, D.E., & Greeno, J.G. (1973). Acquiring cognitive structure by discovery and rule learning. *Journal of Educational Psychology, 64,* 85–97.

Ericsson, K.A., & Simon, H.A. (1984). *Protocol analysis: Verbal reports as data.* Cambridge, MA: MIT Press.

Eylon, B., & Reif, F. (1984). Effects of knowledge organization on task performance. *Cognition and Instruction, 1,* 5–44.

Gagné, R.M., & Brown, L.T. (1961). Some factors in the programming of conceptual learning. *Journal of Experimental Psychology, 62,* 313–321.

Gagné, R.M., & White, R.T. (1978). Memory structures and learning outcomes. *Review of Educational Research, 48,* 187–222.

Heller, J.I., & Reif, F. (1982). Knowledge structure and problem solving in physics. *Educational Psychologist, 17,* 102–127.

Heller, J.I., & Reif, F. (1984). Prescribing effective human problem-solving processes: Problem description in physics. *Cognition and Instruction, 1,* 177–216.

Katona, G. (1940). *Organizing and memorizing.* New York: Columbia University Press.

Langston, I.W., & Watkins, T.B. (1980). *SAT-ACT equivalents* (Research Memorandum 80-5). University of Illinois, Office of School and College Relations.

Larkin, J.H. (1981). Enriching formal knowledge: A model for learning to solve textbook physics problems. In J.R. Anderson (Ed.), *Cognitive skills and their acquisition.* Hillsdale, NJ: Erlbaum.

Mayer, R.E. (1974). Acquisition processes and resilience under varying testing conditions for structurally different problem-solving procedures. *Journal of Educational Psychology, 66,* 644–656.

Mayer, R.E., Stiehl, C.C., & Greeno, J.G. (1975). Acquisition of understanding and skill in relation to subjects' preparation and meaningfulness of instruction. *Journal of Educational Psychology, 67,* 331–350.

Moreira, M.A., & Santos, C.A. (1981). Influence of content organization of students' cognitive structures in thermodynamics. *Journal of Research in Science Teaching, 18,* 525–531.

Novak, J.D., & Gowin, D.B. (1984). *Learning how to learn.* New York: Cambridge University Press.

Perkins, D.N., & Salomon, G. (1989). Are cognitive skills context bound? *Educational Researcher, 18,* 16–25.

Reif, F. (1985). Acquiring an effective understanding of scientific concepts. In L.H.T. West & A.L. Pines (Eds.), *Cognitive structure and conceptual change.* Orlando, FL: Academic.

Royer, J.M. (1986). Designing instruction to produce understanding: An approach based on cognitive theory. In G.D. Phye & T. Andre (Eds.), *Cognitive classroom learning: Understanding, thinking, and problem solving*. San Diego, CA: Academic.

Sherwood, B.A. (1983). Pseudo work and real work. *American Journal of Physics, 51*, 597–602.

Wertheimer, M. (1945). *Productive thinking*. Chicago: University of Chicago Press.

West, L.H.T., Fensham, P.J., & Garrard, J.E. (1985). Describing the cognitive structures of learners. In L.H.T. West & A.L. Pines (Eds.), *Cognitive structure and conceptual change*. Orlando, FL: Academic.

White, R.T. (1985). Interview protocols and dimensions of cognitive structure. In L.H.T. West & A.L. Pines (Eds.), *Cognitive structure and conceptual change*. Orlando, FL: Academic.

APPENDIX

Problems Used in Think-aloud Protocol Sessions

1. A force of 30 Newtons is applied continuously to a box of mass $m_1 = 10$ kg. This box is in contact with a box of mass $m_2 = 20$ kg, and both boxes are on a frictionless surface.
   a) Assuming the boxes remain in contact, find the acceleration of m_2.
   b) What is(are) the magnitude(s) of the force(s) acting on m_2? (See Figure 6.)
2. A mass $m_1 = 5$ kg is on a frictionless surface and connected to a mass $m_2 = 2$ kg by a massless rope which travels over a frictionless pulley as shown. What is the acceleration of m_1? (See Figure 7.)
3. A boy (mass 50 kg), a girl (mass 40 kg), and a block (mass 10 kg) are situated as shown. The boy and the block are on a frictionless surface and free to move, but the girl is standing on a carpet and doesn't move. The boy is holding onto the block by means of a massless, rigid rod (the block is essentially connected to the boy). The girl pulls on the boy with a constant force of 60 Newtons, so that the boy and the block together have an acceleration of 1 meter per second squared. Find the tension in the rod connecting the boy and the block. (Recall that the tension in a massless rod has the same value everywhere in the rod.) (See Figure 8.)

Problems Contained in the Written Exam

Following are the instructions and problems used in the written exam. Problems 1, 3, 5, and 7 are the familiar problems. Problems 2, 4, and 6 are the transfer problems.

1. A block of mass 10 kg is on a frictionless incline (60 degree angle) as shown. Find the acceleration of the block. (See Figure 9.)
2. Blocks 1, 2, 3, 4, and 5 each have a mass of 5 kg, and are connected by ropes of negligible mass. A force of 400 Newtons is applied horizontally on block 5. The blocks are sitting on a frictionless surface.
   a) Find the acceleration of block 5.
   b) Find the tension in the rope between masses 3 and 4. (See Figure 10.)

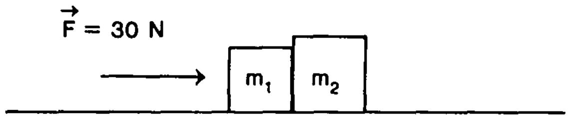


Figure 6.

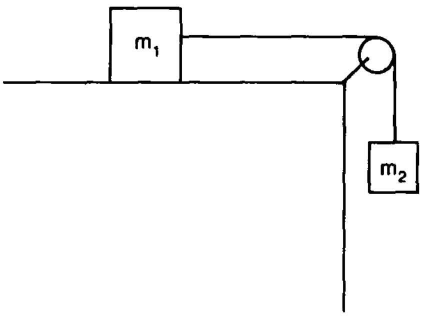


Figure 7.

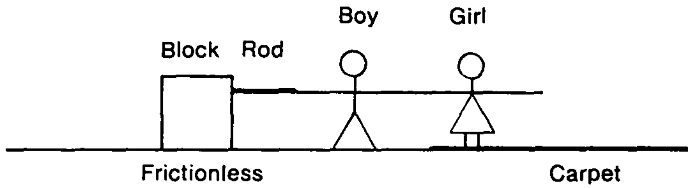


Figure 8.

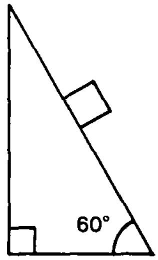


Figure 9.

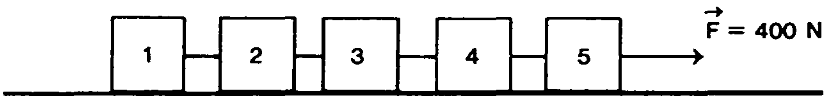


Figure 10.

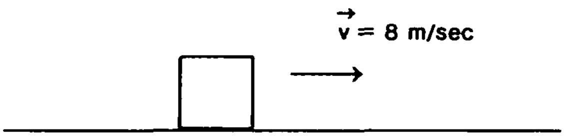


Figure 11.

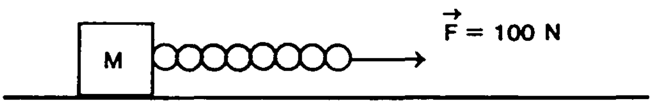


Figure 12.

3. A block of mass 4 kg has an initial velocity of 8 m/sec. There is a frictional force between the block and the surface on which it slides of 10 Newtons. How far will the block slide before it comes to rest? (See Figure 11.)
4. A continuous force $F = 100$ Newtons is applied to a chain which is attached to a block M of mass 30 kg. Each of the eight links in the chain has a mass of 1 kg.
   a) If M is sitting on a frictionless surface, find its acceleration.
   b) If the chain has one link which is weaker than the rest, where would you put that link (how far from the block?) so that M can experience as large an acceleration as possible without the chain breaking? Explain your answer. (See Figure 12.)
5. Shown are masses m_1 (5 kg) and m_2 (2 kg) which are connected by a massless rope which travels over a frictionless and massless pulley. Find the acceleration of m_1. (See Figure 13.)
6. A heavy rope which has a uniformly distributed mass of 14 kg is being pulled upward by a force of 420 Newtons. What is the tension in the rope ¾ of the way along its length? (See Figure 14.)
7. Masses m_1 (4 kg) and m_2 (5 kg) are on a frictionless surface and connected by ropes of negligible mass. What force F is necessary so that m_2 will experience an acceleration of 3 meters per second squared? (See Figure 15.)

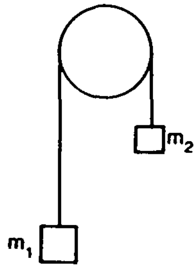


Figure 13.

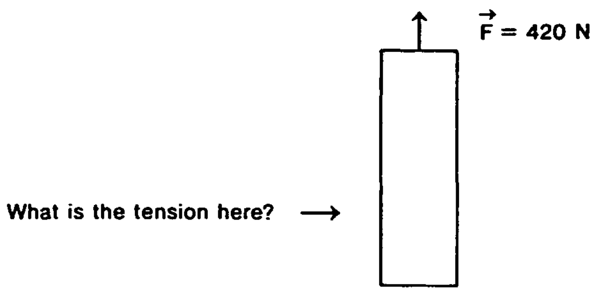


Figure 14.

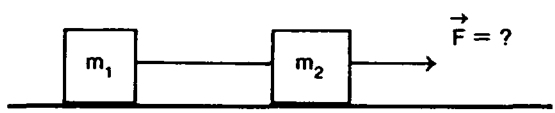


Figure 15.

Selected Codes from the Coding Scheme for the First Protocol Problem

The following codes are those determined to be associated with a subject's understanding of the system concept for protocol Problem 1. The label for each code, such as PS 5 and TD 9e, refers to the code's location in the major categories for each coding scheme. Those major categories are Basic Description (BD), Theoretical Description (TD), Exploratory Analysis (EA), Metacognitive Statements (MS), Problem Solution (PS), and Assessment (A). The codes are grouped below according to the categories for positive and negative indicators discussed in the body of the article. These codes represent approximately 40% of the total number of codes used in the analysis. A complete coding scheme is available from the author.

Positive Indicators

Explicit choice of system:

PS 5 Choice of ____ as a mass to isolate.
a. m_1
b. m_2
c. m_1+m_2

PS 6 Statement that the masses should be combined.

Statement that a choice of system is possible:

PS 4 Explicit statement that a system choice is possible.

Statement that the accelerations of the two masses are the same:

PS 13a State that the acceleration of m_1 equals the acceleration of m_2, or that the acceleration of either mass is the same as that of both together.

Statement that the forces acting on the two masses must be different:

PS 19 State that the force of 30 Newtons is somehow reduced before it gets tom_2, or that only a portion of the 30 Newton force acts on m_2.

PS 21 State that the forces on m_1 and m_2 are different.

Substitution of algebraic symbols or numbers into $\Sigma F = ma$ that are consistent with a given system choice (including stating that given forces are acting on an appropriate system):

TD 5 Assign a value of 30 kg to the total mass (m_1+m_2).
TD 9e Assign a force of 30 Newtons acting on m_1+m_2

PS 8 Assign a force of 20 Newtons acting on m_2

PS 16 Plug in value of F = ____ and m = ____ in $\Sigma F = ma$.
a. 30, 30
e. 20, 20

PS 18 Plug in values of m = ____ and a = ____ in $\Sigma F = ma$.
a. 10, 1
b. 20, 1

Negative Indicators

Statements that the forces acting on the two masses are the same:

PS 20 State that m_1 and m_2 feel the same force, or that the force on m_1 (30 N) is transferred undiminished to m_2.

Statements that the accelerations of the two masses are different:

PS 13b State that accelerations are not the same

Inappropriate substitution of symbols or numbers into $\Sigma F = ma$:

TD 9c and TD 9d Assign a force of 30 Newtons acting on m_2.

PS 8 assign a force of ____ Newtons acting on ____.
a. 20, m_1
g. 10, m_2

PS 16 Plug in values of F = ____ and m = ____ in $\Sigma F = ma$.
b. 30, 20
c. 30, 10
d. 20, 30
f. 20, 10

PS 17 Putting any values other than F = 30 or 20 and m = 10, 20, or 30 into $\Sigma F = ma$.

PS 18 Plug values of m = ____ and a = ____ in $\Sigma F = ma$.
d. 10 anything but 1
e. 20, anything but 1
f. 30, anything but 1

Sample Coding of Think-aloud Protocols

Below are two sample codings of protocols for Problem 1. Codes that contribute to the calculation of the index of system-concept understanding are so indicated (+ or −) and are written out. Other codes describe processes not related to the system concept. (NC = no code assigned).

Subject A:

STATEMENT	CODE	
ok, so...F equals 30 Newtons	TD 9 object not specified	
m_1 equals 10 kg	TD1	
m_2...equals 20 kg	TD3	
ok, let's see...F equals...ma	PS10	
let's see...m equals...m_1 plus m_2	PS6 (Statement that the masses should be combined somehow)	+
let's see...and m_2 is 20 kg	TD3	
m_1 is 10	TD1	
F equals...let's see, assuming the boxes remain in contact, find the acceleration of m_2	BD1	
ok, so let's see...	NC	

I don't think, let's see, m doesn't equal m_1 plus m_2	NC	
F, F equals ma	PS10	
and that m...F, er, a equals F over m	PS2	
let's see, I think...we can...instead of taking them as combined, we can... say that...a force of 30 Newtons acts on m_1, and a force of 30 Newtons also acts on m_2	TD9a TD9c (Assign a force of 30 N acting on m_2)	–
let's see...so, let's see, umm	NC	
F equals m_2a...so F over m_2 equals a	PS10, PS2	
F is 30 Newtons, which is 30 kg per meter squared divided by m_2, which is ...20 kg	PS16b (Substituting $F = 30$ $m = 20$ kg into $F = ma$)	–
so let's see...a equals...three halves meters per second squared	PS1	
what is, or are the magnitude...of the force acting on m_2?	BD1	
F equals...let's see, a force of 30 Newtons acting on that...on m_1...the force acting on m_1...is, let's see...the same thing	TD9a	
so the acceleration would be... acceleration is different	NC	
F over m_1 equals a...that's 30 Newtons...divided by...10 Newtons... equals...not 10 Newtons, 10 kg	PS16c,PS2 (Substituting $F = 30N$ and $m = 10$ kg into $F = ma$)	–
Newtons are...kg meters per second squared	PS14	
kg divide out	PS2	
so a equals 3 meters per second squared	PS 1	
so, let's see...acceleration of m_1 is... 3 meters per second squared...and so, let's see...acceleration of m_2...must also be 3 meters per second squared	PS13a,PS1 (State that the acceleration of m_1 equals the acceleration of ma_2)	+

m_1 is 3 meters per second squared... acceleration	PS1		
m_2 is three halves meters per second squared...and if it was 30	NC		
but, they remain in contact...so m_1... must push m_2	BD7,PS8i		
so let's see...three meters per second squared	NC		
I'm kinda baffled...let's see	MS		
let's see, we got the acceleration	BD2d		
F equals ma	PS10		
F, the force of one...equals m_1a	PS10		
m_1 is 10, times 3 equals 30	PS1		
so...the force of...so the force m_1 exerts on m_2...is 30 meters, is 30 Newtons...if you multiply mass times acceleration	TD9c PS8i	–	(Assign a force of 30 N acting on m_2)
so...the force acting on m_2, is also 30 meters...which gives the original answer of three halves meters per second squared	TD9c PS1	–	(Assign a force of 30 N acting on m_2)
I'm not quite sure what the answer is, but I'd guess that for part a that it would be close...I'd put three meters per second squared as the answer	PS1		

Subject B:

STATEMENT	CODE	
the accleration of m_2 is the same as ...m_1, which is the same as...m_1 plus m2	PS13a (Statement that the acceleration of m_1 is the same as the acceleration of m_2)	+
so...the force is applied to...the two masses, m_1 plus m2	TD9e (Assign a force of 30 N acting on m_1 and m_2)	+
and the, we're looking for the acceleration...of the system	BD3a uo	

so. . .a equals F. . .over mass total. . . which equals. . .30 Newtons divided by . . .m_1 plus m_2. . .which equals. . .30. . . over 30. . .30 equals. . .one meter per second squared. . .acceleration of. . . m_1 and m_2	PS16a, PS1 (Plug in values of F = 30 N and m = 30 kg in F = ma)	+
well, let's see. . .the uh. . .magnitude of the force acting on m_2 is what's left over. . .after acting on m_1	PS19 (State that the force of 30 N is reduced before it gets to m_2)	+
so, let's see. . .so. . .force on two. . . equals force applied minus force on one	PS19 (State that the force of 30 N is reduced before it gets to m2)	+
that doesn't make any sense	MS	
If there's a rope between the two, I can separate the two.	EA	
so I would say. . .that uh. . .the force to accelerate. . .mass two. . .plus the force to accelerate mass one. . .plus force one equals the force to accelerate the whole system.	PS19 (State that the force of 30 N is reduced before it gets to m_2)	+
so. . .it takes, since. . .we know that the accel, ahh, ok	NC	
we know what the acceleration is, so. . .umm. . .force on two. . .equals mass two, 20 kg. . .times a. . .which is equal to one meter per second squared and that makes sense, ok	BD2d, PS18B (Plus in values of m = 20 kg and a = 1 in F = ma) A3	+